MEDIÁLNÍ STUDIA
MEDIA STUDIES
JOURNAL FOR CRITICAL MEDIA INQUIRY

Visualizing COVID-19: an analytical model to understand and compose continuously evolving data visualization projects

Marju Himma-Kadakas, Pille Pruulmann-Vengerfeldt & Signe Ivask

To cite this article:
Himma-Kadakas, M., Pruulmann-Vengerfeldt, P., Ivask, S. (2021).Visualizing COVID-19: an analytical model to understand and compose continuously evolving data visualization projects. *Mediální studia*, 16(1), 65–91.

ISSN 2464-4846
Journal website: https://www.medialnistudia.fsv.cuni.cz/

1/2022

VISUALIZING COVID-19: AN ANALYTICAL MODEL TO UNDERSTAND AND COMPOSE CONTINUOUSLY EVOLVING DATA VISUALIZATION PROJECTS

MARJU HIMMA-KADAKAS, PILLE PRUULMANN-VENGERFELDT & SIGNE IVASK
University of Tartu | Malmö University | Masaryk University

ABSTRACT

The increased demand for information during the Covid-19 pandemic inspired projects to describe the pandemic's progress via data visualization. Critically analyzing the published data visualization projects (DVPs) contributes to establishing a framework that supports both understanding and composing DVPs that evolve over time. Drawing upon constructed grounded theory, we develop an analytical model for creating DVPs in a journalistic or public communication context. For our analysis, we selected Covid-19 public service media DVPs in the United Kingdom, Norway, Sweden and Estonia as well as DVPs created by global and local data activists. The analysis of these examples provides an understanding of (1) the implied agency standing of the authors of the visualizations, (2) the kinds of editorial layer (data, visual representation, annotation or interactivity) that inform the creation process and (3) what newsrooms and data visualizers can learn from this practice to create understandable, meaningful and engaging DVPs of (critical) events that evolve over an extended period. Our model supports data visualization practitioners in making informed choices when creating data stories.

Keywords: data visualization ▪ data journalism ▪ data activism ▪ data literacy ▪ digital journalism ▪ Covid-19

1. INTRODUCTION

Continuous data reporting on Covid-19 became part of the fight against the pandemic, and 'Flatten the curve' became the mantra of various early initiatives to control or combat the coronavirus. Questions remained, however: Where is the curve? What does it consist of? Moreover, will the curve predict the outcome for us? For our

country? What does the curve even mean? Countless data journalists, public officials and data activists put effort into illustrating the curve.

The visualizations predicted the bell-shaped curve to become flat again after the pandemic evolved and peaked, but retrospective analysis showed the shortcomings of such visuals. For example, they tended to be short-sighted, inaccurate and unconnected to scientific data (Pentzold et al., 2021). Critical voices argued that the inaccuracy of data visualizations merely contributed to an 'infodemic' (Mooney & Juhász, 2020). The aims of Covid-19 data visualizations have changed over time from awareness-raising to calls for responsible behavior, from fear-mongering to fostering public discussions. The intended readers, their skills and knowledge at understanding and interpreting the data visualization have been challenged a lot. Covid-19 and many similar crises (e.g. the war in Ukraine or the heatwaves emerging over a long period) offer case studies to inform a variety of global and national journalistic data visualizations of pandemic development and, as such, an excellent opportunity to learn about visualizations as part of continually evolving DVPs.

When the curve did not flatten as initially predicted, news organizations and public health institutions were tasked with creating information visualizations that acknowledged the uncertainty caused by the evolving situation, were continuously updated and were intended to be functional for diverse target groups. This made us ask how research could contribute to establishing data visualization projects (DVPs) that need to continue over time and contribute to disseminating meaningful data-based knowledge.

This paper investigates Covid-19–related DVPs from public service media websites, international data activist initiatives and public health authorities in Estonia, Norway, Sweden and the United Kingdom. We use the term *data visualization project* to embrace all the diverse types, from stories illustrated by data visualizations to DVPs supplemented with some text. By analyzing the Covid-19 visualizations' authorial agency, editorial layers and communicative functions, we establish a heuristic model that supports the creation of continually evolving DVPs. We take inspiration from Hullman and Diakopoulus (2011), who argue that investigating data visualizations on the basis on the evidence in their visuals allows us to understand some of the editorial judgements underlying their rhetorical techniques. In developing our model, we drew upon the theoretical understandings of several journalistic theories in conducting reflexive iterations of qualitative analysis.

The specific case of Covid-19 DVPs allows us to analyze data visualizations beyond the traditional data stories in which visuals support journalistic storytelling; in this case, the visualizations were available and used by audiences without the narratives that would contextualize them. Our analysis takes these 'context-poor' visuals as the starting point for our model building. Our heuristic model is usable in journalism and many other fields in which creating continually evolving DVPs is necessary.

Our qualitative analysis based on reflexive iterations—repeatedly visiting the data and theoretical approaches—enabled us to identify three dimensions that constitute

the basis of the heuristic model. **The analysis of examples of various international DVPs shows (1) that variations in the *agency* of diverse actors explain their positions and role-taking, (2) that the *editorial layer* reflects a balance between automation, time budgets and journalistic freedom and (3) that the overall *communicative function* supports the receiver in interpreting a visualization**. By looking at the complexity of the data stories, we learn from continuing Covid-19 reporting practices how to create understandable, meaningful and engaging DVPs of (critical) events that develop over an extended period.

2. VISUALIZING DATA IN TIMES OF CRISES

With the advent of data journalism, the need to visualize vast amounts of data and an increased attention to automation have become journalistic norms. In using data in stories, however, journalists may also choose pre-packaged news angles that interest the audience, and the choices of interactive data presentation may be relatively limited to fulfil their paternalistic perspective (Appelgren, 2018). In the case of the Covid-19 pandemic, the actors who engaged in visualization were journalists, public health authorities and data activists who rushed to fill the information vacuum and satisfy doomscrolling audiences (Pruulmann-Vengerfeldt, 2020). The daily reporting of pandemic data forced many news organizations to establish continually (self-) updating forms of DVP. The various actors who contributed to their data stories had diverse interests and focuses, resulting in a diversity of visualization approaches (Pelizza, 2020).

Due to the social constructionist nature of data visualizations, the notion of objectivity is inapplicable in data journalism (Tong & Zuo, 2021). Masullo et al. (2021) detected a *crisis coverage gap*, showing that the media did not consistently deliver what the public really needed during the Covid-19 pandemic, indicating a misalignment of journalists' roles and audience expectations. Journalists' agency significantly stems from their role perception, and that perception informs their published stories as is shown in several works on journalistic role performance (e.g., Hanitzsch et al., 2011, 2019; Mellado, 2015). Nevertheless, as an important starting point of any communication act, the significance of role performance in editorial processes is often underestimated or not discussed or described.

Overall, data journalism expands the possibilities for informative storytelling at the intersection of the fields of informatics (e.g., programmers and data analysts) and communication (e.g., journalists). Data/computational journalism is rooted in professional journalism and open-source culture (Coddington, 2015), the latter of which somewhat relates to the recently emerging phenomenon of data activism. On the one hand, Wu et al. (2019b) show that the technological firms that provide technical automation and data-related services to newsrooms distance themselves from data storytelling, responsibilities and consequences. On the other, Örnebring and Ferrer-Conill (2016) look at this process from the newsroom's perspective and

conclude that the outsourcing of information and data (visualization) services means that potential skills and competencies in the newsroom remain underdeveloped. Additionally, Örnebring and Ferrer-Conill (2016, p. 4) show that information and data services are outsourced by news agencies mainly for automated financial data processing. They also point out that newsrooms have to deal with increasing amounts of data while coping with growing pressure to make news production more efficient. At the same time, news workers see machines as cocreators of news alongside humans in the automated newsroom (Primo & Sago, 2014). Automated journalism may mean anything from machine aggregation of content to data scraping and auto-publication (Wu et al., 2019a). The diverse understanding of automatization also sets conditions for establishing newsroom routines, where continually evolving DVPs demand a balanced and carefully considered allocation of machine and human labor.

Outsourcing data services for specific stories is uncommon, and economic pressures support the use and publishing of already 'packaged' data, meaning that almost no effort is invested in the innovation of data visualization in house (Örnebring & Ferrer-Conill, 2016). As Örnebring and Ferrer-Conill discuss, outsourcing also means that outside actors (freelancers) determine the frame of the news and may even neglect some journalistic norms. Our analysis allows us to challenge this claim to a certain extent. Although the data underlying Covid-19 data 'packages' was the same in several of our examples, we identified a diversity of visualization possibilities due to the differences in agency, editorial layer and communicative function attributed to the visualization.

Matters related to the source, limitations and data interpretation are essential to data visualization's transparency, so it is essential that skilled, trustworthy actors make the visualizations. Kõuts-Klemm (2019, p. 311) has argued that journalists who possess more in-depth experience of sociological methods or statistical concepts use that experience in evaluating the quality of information more often than those who do not. Tong and Zuo (2021) concur and argue that data processing is an arena of subjectivity and decision-making. Nevertheless, most data literacy studies conducted among journalists do not discuss their ability to assess the media and/or data literacy of the receiver (the audience) (Kennedy et al., 2016). As the Covid-19 pandemic started, several outside actors assumed the role of trustworthy, authoritative actors (e.g., Johns Hopkins University), reducing the need to address the (possible) lack of journalistic skills in data processing.

The inaccuracy of data visualizations was criticized during the pandemic for contributing to the infodemic (Mooney & Juhász, 2020). Differences between public authorities that collected and reported Covid-19–related data meant that apples were often compared with oranges in visualizations, contributing to mistrust and confusion. Journalists who use pre-framed data without analyzing its limitations are just one problematic side of data storytelling along with questions regarding the transparency and credibility of the data and the source.

Presenting the data source may function to establish trust and increase the overall transparency of the news and its production process, but a profusion of source links can also create an illusion of trustworthiness (Phillips, 2010; Rupar, 2006). Audiences appreciate transparency disclosures, such as hyperlinks to the original content or other sites as well as the acknowledgement of errors in reporting (Karlsson & Clerwall, 2018, p. 1927–1928). Therefore, linking directly to the initial source of the data in a data set, explaining the methodology and disclosing possible controversies in the data and visualization increase a data story's credibility (Kennedy et al., 2016; Potter, 2004). Awareness of the value of such practices helps journalists and other data mediators make choices that support the audience's data literacy.

Instead of normatively evaluating Covid-19 visualizations, our approach pinpoints how three distinct processual elements—agency, editorial layer and communicative function—provide a heuristic for data visualizations.

3. THEORETICAL APPROACHES TO DATA VISUALIZATION

We used the inductive approach of reflexive iteration, which involved visiting and revisiting the collected DVPs and comparing the reflections to the theoretical concepts and extant scholarly knowledge. This led to refined insights into the domains of the model. For these insights to make sense, it was essential that links be established to existing theoretical perspectives. Due to the situation's complexity, we drew upon multiple theoretical perspectives in building the heuristic model to determine agency, editorial layers and the communicative functions of the DVPs.

3.1. Agency

In this paper, we rely on existing research on journalistic roles to understand the agency of authors, because, while diverse social actors made Covid-19 visualizations, their positionality resembled that of a journalistic role even if they were not journalists. Previous studies have used diverse methods (Hanitzsch et al., 2011; Hanitzsch & Mellado, 2011; Hanitzsch & Vos, 2018; Mellado, 2015) and identified journalistic roles in three domains, each of which contains several specific roles.

The first, the *power relation* domain, includes the watchdog role, in which the journalist monitors actors and institutions, as well as the loyal facilitator role, in which the journalist cooperates with those in power or expresses blind loyalty to institutional elites and the nation-state. The second, the *journalistic voice* domain, embraces the disseminator role, in which journalists distance themselves from the content, and the interventionist role, which involves journalists' interpretations, opinions or other elements. The third, the *audience approach* domain, contains three professional roles: the civic role (audiences are seen as citizens), service role (audiences as consumers) and infotainment role (audiences as spectators).

In making the Covid-19 data stories, the authors—regardless of who they

were—took on roles identified in previous research, and thus their agency positions may be identified in the visualizations they produced.

3.2. Editorial layer

The concept of editorial layers was introduced by Hullman and Diakopoulus (2011, p. 2233). In their understanding, the concept embraces four layers—data, visual representation, textual annotation and interactivity—which allow one to look at visualizations to understand their authors' rhetorical choices. Editorial layers are strongly linked to the elements of journalistic storytelling, such as news value and framing, and we endorse the argument that they reveal the choices made by the maker of the visualization. Therefore, the concept provides a frame that lets us trace the choices made in creating visualizations. In journalism, some of these choices have been extensively studied and outlined, so we build our reading of the editorial layer on existing theories of news values, and we extend them to the discussion of automatization, a key element in building data visualizations of continually updated data.

Data visualization requires editorial decision-making that is largely based on **news values**. Many of the news values identified by Galtung and Ruge (1965) and Harcup and O'Neill (2001, 2017) are present in DVPs. The news values selection in Galtung and Ruge's (1965) chain of news communication occurs in the three domains of the chain: media perception, selection distortion and media image. This is an essential part of the editorial layer, as that which is accentuated is evident in the outcome (e.g. the story or visualization) serving the communicative function. Not to delve into the details of all the news values proposed by several authors, we underline some that are significant in Covid-19 DVPs.

Personification as a news value means setting individual actions or people in the foreground of reporting and it is relevant in identifying the event, as people can feel empathy for other persons more easily (Galtung & Ruge, 1965; Harcup & O'Neill, 2017). In data visualizations, however, personification in graphs and tables must be achieved using other means, but which means? While the audio-visual element itself can function as a news value (Harcup & O'Neill, 2017), the question arises of whether this news value has a central or secondary position in DVPs. Frequency and continuity as news values refer to events' continuous or regular occurrence (Galtung & Ruge, 1965; Harcup & O'Neill, 2001, 2017). In continually evolving DVPs, the pace of updates (which may diverge from the overall news cycle) sets the frequency factor. In evolving DVPs, the frequency factor can be closely related to the news values of exclusivity (exclusive access to information) and conflict (confrontation between counterparts), as the information may be available only to the DVP author, and every update may conflict with the previous data presentation (Harcup & O'Neill, 2017). Conflict is sometimes united with the value of bad news to visualize unexpected, tragic and morbid events as constituents of the group of news values arousing negativity. The

Covid-19 pandemic by its nature meets the criterion of bad news. However, as the pandemic evolved, the criterion of good news began to gain relevance. Therefore, the accentuation of these two values played a role in evolving DVPs. The Covid-19 crisis was simultaneously geographically close and globally distant, incorporating geographical proximity as a news value. The extent of accentuating proximity depends on the author's agenda, which, conversely, depends on ethnocentrism or cultural proximity (Galtung & Ruge, 1965). The crisis involved select countries and elite people; it was personal, meaningful and relevant, but in diverse ways depending on the DVPs' position and the agenda of the outreach.

When an event evolves over a long time (as did the Covid-19 pandemic), DVPs must make the process of continual updating routine; in other words, they must rely on **automatization**. Tuchman's (1973) questions on routinising the unexpected ask how news organizations process unexpected events and how news workers decrease the variability of events in the news. Covid-19 coverage would generally be classified as developing news, but, because developing news, according to Tuchman (1973, 117), must be unscheduled, the daily updated data visualizations count as continuing news, as they imply the existence of prescheduled change. Regional and international Covid-19 data publication was prescheduled by the data collectors, and journalists, public authorities and data activists developed routines for updating their data stories. However, the Covid-19 DVPs did not merely persist but also evolved over time. Therefore, we prefer the term *evolving DVPs*, as it describes both the continuity and evolving nature of the projects.

The automation of journalistic processes can have the objective of replacing journalistic work (e.g., Carlson, 2015) or can free journalists' time for reporting (Wu et al., 2019b). Tuchman (1973) sees the routinisation of work processes and resource allocation as part of managerial decision-making in the newsroom. Automating and outsourcing some activities in evolving news reporting enables journalists to engage in storytelling or in covering other news (Coddington, 2015).

Caswell and Dörr (2018) highlight the dichotomy between automatization and the editorial challenges of expressing narrative structure in a grammatically and structurally coherent manner in automatically generated news. Therefore, the dimensions of automatization and author autonomy are interdependent. Drawing from previous conceptualizations, we bring together the intersecting axes of automatization, time and author autonomy under the editorial layer.

Regardless of the agency behind editorial decisions, visualizations are also related to resource allocation. The editorial layer reveals the homogeneity or diversity of news values and determines the process of updating information. Manual updating of data demands human labour and is time consuming. At the same time, it gives journalists and other actors more autonomy to adjust the story (e.g., by covering the topic in light of the event's current focus, by changing news values or by simultaneously covering different data sources). Since the time spent on a story is an expensive commodity and journalists' or data activists' input can be more valuable elsewhere, it

is cost effective to automate data stories. However, when left unattended, the visualization may become detached from the actual conditions of the situation. Therefore, we see that the interrelations between time, author autonomy and automatization are critical dimensions of the editorial layer and are directly related to the diversity or homogeneity in the chosen news values.

3.3. Communicative function

We draw from the concepts of visual framing and engagement to define communicative function, which, according to Weber's (2020) characteristics of a data story, embraces visual framing and engagement specific to data storytelling (Engebretsen et al., 2018; Kennedy et al., 2016). As several aspects of the communicative function are closely related to the editorial layer, our understanding also draws upon Hullman and Diakopoulos's (2011) narrative visualization characteristics, including data provenance strategies, representation of uncertainties and individualization.

Data visualizations derive their communicative functions from **visual framing**: the selected graph types, colors, symbols, icons, etc. As choices must be made to create the visualization, it can be argued that decision-making bias is unavoidable, and choices subsequently influence how the visualization is perceived and interpreted. Data constitute the foundation of DVPs, and, in our case, it is vital that data are regularly updated and represent, first and foremost, the number of people diagnosed with Covid-19 and the number who died or recovered from the disease. Both purpose and annotation both include keywords indicating the communicative intent of the producers (e.g., to tell, to explain, to argue visually) (Weber, 2020). Hullmann and Diakopoulos (2011) limit *annotation* to explicitly labelling parts of the visualization or text to represent uncertainty or expressions of doubt regarding potential conclusions. The textual-visual relationship links the visualizations to the text and is best represented in 'scrollytelling' techniques (i.e. a type of story in which the continuous scrolling of the page creates the story experience and engages the user with various media). Visual design frames the emotions, morbidity and stereotypes explicitly or implicitly represented in the DVP. The information structure has several forms, including author-driven/linear/explanatory and reader-driven/nonlinear/exploratory narratives. By selecting some and neglecting others of these characteristics, DVP authors frame the message conveyed to the audience.

Information literacy requires the knowledge and ability to use the technological features of data storytelling (Engebretsen et al., 2018), which is inseparable from what attracts the user and can be summed up as **engagement**. Kennedy et al. (2016) identify six factors that affect engagement, which we incorporated in the heuristic model through the contextual interpretation of the Covid-19–related data visualizations:

1. Subject matter: determines the user's possible interest in the subject of the story. In the Covid-19 case, timely actual coverage could be part of audiences' coping strategies.

2. Source/media location: refers to the publisher (e.g. news organization). For the heuristic analysis, we looked at Covid-19 data visualizations from diverse agents.
3. Beliefs and opinions: visualizations may intensify the user's beliefs and/or opinions on the information. Data visualizations have the power to confirm or contest users' beliefs and opinions on the spread of Covid-19.
4. Time: visualizations offer the possibility of spending time through interactive usage or may require time to understand. Covid-19 data provide complex data sets that may take time to interpret and understand.
5. Emotions: when users have engaged with the subject matter and given time to the data story, the visualization may inspire strong emotional responses, which are evoked by visual design elements (e.g., signs, colors, icons).
6. Confidence and skills: users need to feel that they have the necessary skills to decode visualizations. For users to understand complex Covid-19 DVPs, the visualizations required an intuitive user interface and guidelines to help viewers interpret and use the data.

To understand the audience's perspective on data stories, it is essential to consider ideas of data and information literacy. *Data literacy* describes the capacity to use and interpret data, and it emerges in the overlapping areas of statistical literacy, information literacy and technical skills in working with data (Gray et al., 2018, p. 3). The data and information literacy aspects of Covid-19 DVPs can be interpreted using Potter's (2004) media literacy axioms; specifically, the *interpretation* axiom (how people make sense of the DVP), *shared meaning* axiom (matching meanings to designated symbols, i.e. word, graphic or photo), *power* axiom (awareness of the media functions in order to be empowered by the knowledge) and *purpose* axiom (users' empowerment to shift control from the media functionality to themselves). These aspects—along with questions related to agency and (journalistic) roles, news values, visual framing, the engagement factors of data stories and the editorial layer relating to time, automation and autonomy—constituted the starting point of data investigation stories about Covid-19.

4. METHOD

As scholars in the fields of audience and journalism studies, we noticed specific patterns while exploring data visualizations of Covid-19. First, we identified agencies (similar to the journalistic role concept of Hanitzsch et al., 2011; Hanitzsch & Mellado, 2011; Hanitzsch & Vos, 2018; Mellado, 2015); these were also visible when the authors of the visualizations were not journalists. Second, we could see traces of the editorial decision-making processes (editorial layer) in the evolving visualizations. Several factors and actions influence the decision-making of what becomes news. We drew on news values (Galtung & Ruge, 1965; Harcup & O'Neill, 2017), as they are considered the primary determinant in 'becoming news', but combined them with

questions about automatization, as the allocation of resources related to data processing, time and author autonomy varied significantly across various actors. Third, we observed a series of patterns in what we call communicative function, which constitutes the meeting place of expectations regarding audience literacy levels, engagement with visualizations and framing. Next, we introduce the three key clusters of patterns that form the three pillars of our theoretical structure for the proposed heuristic model.

There are many Covid-19 visualization platforms all over the world, making it difficult, if not impossible, to collect a comprehensive sample. Pászto et al. (2020) attempted to systematically categories and critically evaluate and map sources of information and data, but, as our aim was to understand the various actors and their agency roles in the data visualizations, we made a targeted selection and included in our analysis data from three types of source: official sources (such as health boards), public service media (in Great Britain, Sweden, Norway and Estonia) and global data initiatives' online pages based in the US (N=11; see Table 1). The selection was guided by the objective of including visualizations from small transitional democracy media environments (e.g., Estonia), medium-size advanced democracy media environments (e.g. Sweden and Norway) and significant, established democracy media environments (e.g. the UK) as well as sources that positioned themselves as global (e.g. Estonian and US cases). Most of the online pages in our sample were in the local national language except in the case of Estonian Public Broadcasting, whose overview of Covid-19 was in English, as this was the only page that used visualizations systematically and consistently.

Table 1. Selection of data visualizations used in the iterative analysis

Country	Public Service Media	Public health authority	Data initiatives
Estonia	Estonian Public Broadcasting (ERR) https://news.err.ee/1061575/coronavirus-in-estonia-all-you-need-to-know	Health Board of Estonia* https://www.terviseamet.ee/et	Koroonakaart.ee www.koroonakaart.ee
United Kingdom	BBC https://www.bbc.com/news/world-51235105	Department of Health and Social Care https://coronavirus.data.gov.uk/	
Norway	Norwegian Public Broadcasting (NRK) https://www.nrk.no/korona/	Norwegian Institute of Public health (FHI) https://www.fhi.no/sv/smittsomme-sykdommer/corona/dags--og-ukerapporter/dags--og-ukerapporter-om-koronavirus/	

Country	Public Service Media	Public health authority	Data initiatives
Sweden	Swedish Public Broadcasting: (SVT) https://www.svt.se/datajournalistik/har-sprider-sig-coronaviruset/	The Public Health Agency of Sweden (FHM) https://www.folkhalsomyndigheten.se/smittskydd-beredskap/utbrott/aktuella-utbrott/covid-19/	
Global/US			The Johns Hopkins University (JHU) https://coronavirus.jhu.edu/map.html Worldometer https://www.worldometers.info/coronavirus/

* Initially Estonian Health Board did not have their own visualisations and referred to the activist initiative site Koroonakaart.ee. In the second wave of the pandemic, the Health Board adapted their own visualisation which looked very similar to the activists' visualisation.

We gathered data by downloading full-length screen captures of data visualizations that mediated Covid-19 information on May 15, July 15 and September 1, 2020. The time periods were selected because May was three months after the pandemic started and was when various DVPs were gaining popularity. We selected July because summertime in the Northern hemisphere usually brings a decline in viral diseases. September was selected as the third time-point, as it could show potential changes and developments in the DVPs. The pages consisted of information on Covid-19 in written form and graphs. We looked at the information's overall presentation and the possible aim or focus of the presented data. As screen captures were made, all the authors had the same data to analyze and code.

We developed our model through reflexive iterations, in which the categories emerge in the process of going back and forth between empirical material and theoretical readings. In qualitative analysis, this method relies on the inductive approach, in which the categories of analysis emerge from the data (Neale, 2016; Srivastava & Hopwood, 2009). Reflexive iteration is a systematic, repetitive and recursive process in which the researcher repeatedly asks questions that serve as the framework for the data analysis; this task is carried out in precisely the same manner each time and executed multiple times (Miles & Huberman, 1994). There are several frameworks for iterative coding, but they are all based on questions for reflexive inquiry. We combined the reflexive questions of two iterative coding frameworks (Srivastava & Hopwood, 2009) with the three dimensions of our operational model derived from the scholarly literature. Personification as a news value contributed to understanding the means by which DVPs can create empathy and a level of personal identification with data.

For iterative coding, we visited and revisited the DVPs with, altogether, 49 questions under the following categories (presented with example questions):

- editorial decision-making on automatization, autonomy and news values (How is cultural proximity/relevance/consonance/unexpectedness, etc. expressed? Are data updated regularly and automatically, and is this explained to the user? What is the balance between automated data presentation and journalistic storytelling?)
- topical framing (What does the storytelling of this DVP emphasize? What tone does the composition of the story express?)
- visual framing (What is the textual-visual content relationship? How are animations, multimodality, colors and the like presented?)
- practical and technical usability (Is numeric information explained to the user? Are technical guidelines (and what sort) given to the user?)
- information literacy (What kinds of technical and mathematical/statistical skill and knowledge are expected of the user?)

The domains were then augmented with categories from theoretical perspectives. In several cases, these studies take the audience's perspective (see Engebretsen et al., 2018; Gray et al., 2018; Kennedy et al., 2016) or that of data journalists (see Kõuts-Klemm, 2019; Weber, 2020; van Wisten, 2020), but, in our discussions, we turned to these categories to identify traces of decisions apparent in texts. We also employed news values, which are usually not discussed in studies on data storytelling. The present study aims to fill this gap by using the conventional news values approach (Galtung & Ruge, 1965; Harcup & O'Neill, 2001, 2017), as it has a strong role in explaining the framings of both journalistic and data activists' data stories. In the iterative analysis, observations from the websites and the literature reading were systematized under three main domains: agency of the author of the visualizations, editorial layer and communicative function (Table 2).

Table 2. Domains and categories for analyzing data visualizations

Domain	Category group and reference	Categories
Agency	Journalistic roles (Hanitzsch et al. 2011; Hanitzsch and Vos 2018; Mellado 2015)	Agency represented through roles (disseminator, interventionist, investigator, loyal facilitator, watchdog, infotainment, civic service)
Editorial layer	News values (Harcup and O'Neill 2001; 2017; Galtung and Ruge 1965; Pelizza 2020)	News values* Communicative function
	Automatisation (Tuchman 1973)	Manual / semi-automated / automated updating

Domain	Category group and reference	Categories
Communicative function	Data literacy (Kõuts-Klemm 2019; Gray et al. 2018; Potter 2004)	Visual-numeric literacy (knowledge of mathematics, statistical literacy, understanding of visualisations) Source transparency (source/media location), methodological transparency, synthesis, and abstraction. Media literacy axioms.
	Visual framing, and storytelling (Weber et al. 2018; Potter 2004)	Grouping and textual-visual relationship, visual design (interactivity, animations, multimediality, etc.), (visual) data framing, the structure of information, technical usability.
	Data-vis engagement (Engebretsen et al. 2018; Kennedy et al. 2018)	The subject matter (beliefs, and opinions, time, emotions, confidence, and skills)

* Complete list on news values follows the news values presented by (Harcup and O'Neill 2001; 2017; Galtung and Ruge 1965)

5. FINDINGS

5.1. Agency in data visualizations

Although we investigated our sample DVPs at three points in time, we saw little change. Only visual framing showed moderate change, which we explain in more detail in the relevant section. However, we begin by describing the findings on agency.

Any DVP is an interpretation that requires taking a position—an *agency*. The authors' agency in our study samples exhibited various roles that had characteristics in common with the journalistic roles outlined by Hanitzsch et al. (2011) and Mellado (2015). Iterative coding distributed the data visualizations into agency categories of *advocacy*, *power relations* and *audience perspective*. Agency, therefore, reflects roles that are somewhat similar in their dimensions to journalistic roles. Agency influences the editorial layer by establishing priorities among selected news values and assigning resources between automation, time and autonomy; it find its outlet in the communicative function, where data literacy, engagement and visual framing are found.

Public service media (PSM) Covid-19 DVPs expressed diverse journalistic roles and formed two groups diverging from the agency domain. The first group, NRK (*Norsk rikskringkasting*) and SVT (*Sveriges Television*), focused purely on visualizing data, and the textual narrative is almost absent (Figure 1). Thus, they fall under the category of the advocacy/disseminator role. The second group, BBC (*British Broadcasting Corporation*) and ERR (*Eesti Rahvusringhääling*), had pages with a much stronger reliance on journalistic techniques, and their data elements take a backseat, serving as illustrations to the main story. Although they also take an advocacy position, the BBC and ERR play a more interventionist role, due mainly to preserving a robust journalistic presence in the textual narrative.

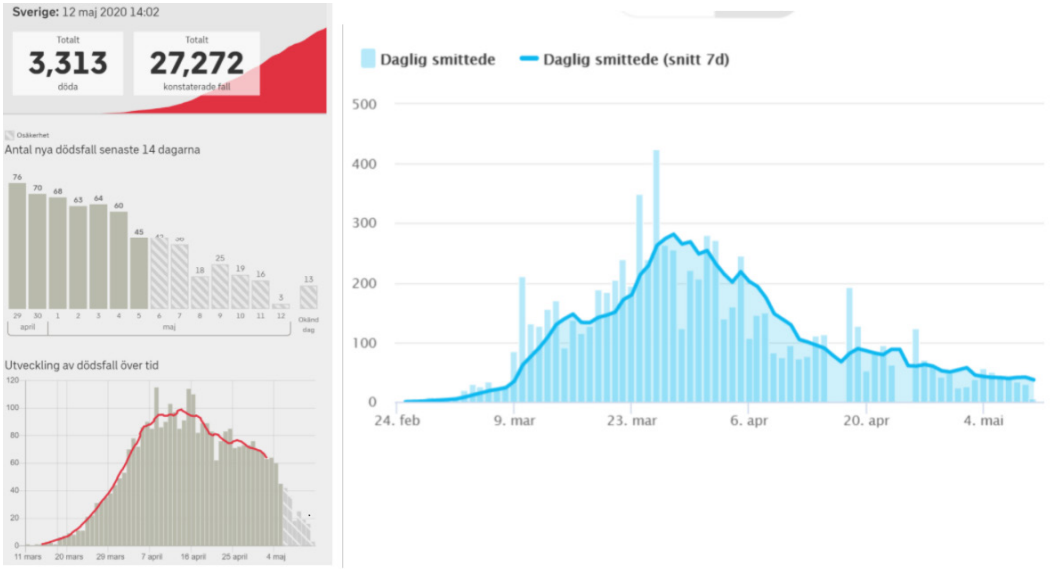

Figure 1. Example of SVT's (left) visualisation of death cases, identified cases, no. of deaths in the last 14 days, change of death cases over time. NRK's (above) bar graph shows the no. of infected, and no. of infected over the last 7 days. (May 12, 2020)

Assessing the content of the PSM DVPs made evident the service and loyal facilitator roles. All the PSM pages were based on official Covid-19 statistics provided by governmental administrations, as this data source is most reliable in these countries. Advice is given in the data visualizations, and the texts follow the official pandemic guidance and do not contest or criticize the guidelines.

Agency in administrative DVPs is strongly influenced by their function, purpose and communicative aims while agency in the data activists' DVPs is expressed in an overview of data without intervention. Koroonakaart.ee and Worldometer provided DVPs without a clear statement of purpose regarding at whom they were aimed or why they were deemed necessary. JHU's (*Johns Hopkins University*) clear interventionist agency can be determined by its statement that the information is 'to help the public, policymakers and healthcare professionals worldwide'. Judging from the targeting of the audience and the presentation of the information, we conclude that public authorities' and data activists' DVPs express the neutral disseminator's role without elements of intervention. The exception is JHU, which states the aim of providing help to the public and fulfils this aim by explaining various media literacy abilities to users.

The agency domain becomes explicit through how it works in conjunction with the editorial layer and communicative function. The DVPs' authors do not actively express agency as is evident from the sample of public authorities' and data activists' visualizations. By simply providing visualizations, they take a detached, disseminator

position, but, as our iterative coding in the editorial layer and communicative function domains shows, adopting this role has its shortcomings.

5.2. Editorial layer

Journalistic and data activist DVPs differ in the news values they adopt. News values help to interest the audience in the story, which makes the selection of news values a central element in the editorial layer. The journalistic DVPs relied on negativity to catch the audience's attention, often representing bad news by the use of contrasting, stereotypical red and black colors to designate deaths or infected cases. By contrast, the Koroonakaart.ee, SVT, and NRK DVPs chose colors that were neither stereotypically gendered nor connotatively loaded (e.g., blue-white) or switched to them over time.

The news value of personification helps to explain the elements by which DVPs create empathy and a level of personal identification with data. Personification, which often appears in company with several other news values, was explicitly present in DVPs that had greater autonomy to change their focus and framing and that must have required more human labor to update. The ERR and BBC DVPs provide examples that strongly emphasize conveying the message by using photos of real people as well as textual supplements that explain the data or emphasize their human impact. However, in the DVPs whose updating was fully automated with no changes in the news values, focus or visual framing, personification was absent as were other news values (such as proximity, meaningfulness and relevance) that often accompany it. Hence, the more automated the DVP, the fewer news values and personification factors were included. Figure 2 provides an example of personification in which ERR included a picture of police officers demonstrating the 2+2 rule. The picture was attached to the story in the early months of the pandemic and remained part of the DVP.

Public spaces

- Restriction change: From June 19, the **2+2** restriction is a recommendation and not a rule.
- Restriction change: **Public events**: 1,500 spectators can attend indoor events and 2,000 at outdoor events from July 15. But capacity must be limited to 50 percent at indoor events
- Restriction change: **Public meetings** can be held from May 18.
- Restriction change: **Playground**s and **sports areas**, such as **outdoor gyms**, will be reopened from May 2 but must be disinfected every four hours.
- Restriction change: Tallinn will reopen **playgrounds** from May 11.
- Restriction change: Tallinn will reopen **public beaches** on June 1.
- RMK is asking **hikers** to avoid popular trails to limit the spread of the virus.

The PPA demonstrate the 2+2 rule in force during the emergency situation. Source: PPA

Figure 2: Example of personification in ERR's DVP: DVP contains text that is manually updated and the picture of the police demonstrating 2+2 rule, personalising the DVP (Screenshot from September 1, 2020)

SVT's DVP, however, evolved over the examined time, and the news value of proximity became more evident, as users could explore the spread of the virus at the regional level. This shows that the visualizations were re-evaluated in the editorial

layer, and changes were made even though most of the story remained automated; we discuss the changes of focus in the DVP below.

Audio-visual elements or, more broadly, multimodality as a news value was emphasized in various ways. All the PSM DVPs emphasized the value of interactive data visualizations as supplements to help the user to understand the situation. However, the interactivity was limited mostly to animation in the presentation of the data and the ability to highlight by selecting one country or a specific variable. This made the message of the PSM visualizations rather unidirectional, aligning with the paternalistic perspective as described by Appelgren (2018). In comparison, independent DVPs, such as Koroonakaart.ee and Worldometer, also emphasized the importance of visualizations but with the purpose of offering an overview and input for other stakeholders to use their visualizations. In addition to graphics, multimodality was present in the ERR and BBC visualizations in the form of videos and photos, which enabled newsrooms to update the DVP and shift its focus in response to journalistically timely emerging topics. This was the main distinction from the other DVPs in our selection, whose purpose was to provide up-to-date numerical Covid-19 data.

Regarding the frequency or continuity of the DVPs, most did not disclose the time of the update at the very beginning. Nevertheless, the update time was clarified according to the data sources' updates. Explaining the regularity of updates clarified the frequency news factor for users, and the practice was present in all the DVPs. The analysis of news values enables us to highlight the relationship between automatization, author autonomy and news values. This is graphically depicted in Figure 3 and explained below.

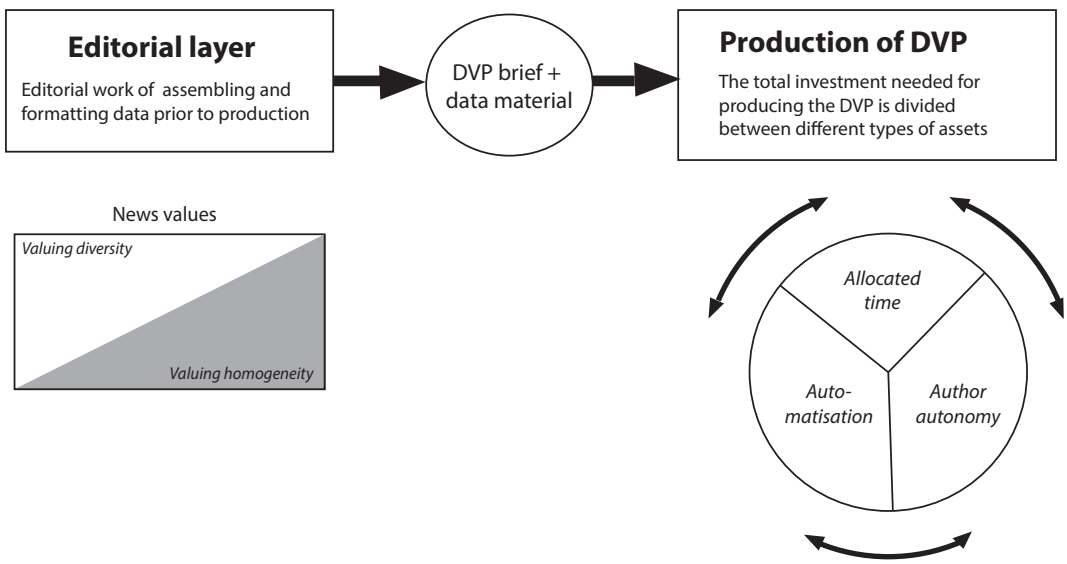

Figure 3. The editorial layer of continually evolving data visualizations

The highly automated updating of DVPs allowed less author autonomy as seen in the static focus and framing of the visualizations. By contrast, greater author autonomy was seen in DVPs that were less automated and more dynamic in changing their focus and framing. The autonomy was explicit in the diversity of news values and in the changes in focus and news values that characterized those DVPs.

In public authorities' and data activists' visualizations, the data in graphs were updated automatically. The text accompanying the graphs was kept to a minimum to minimize the necessity of manual updating. As a result, we deduce that the time spent on the daily updating of the visualization was minimal, giving the authors the freedom to engage in other activities or, in the case of public authorities, to fulfil their primary duties.

The BBC and ERR DVPs with embedded visualizations (including from third parties) seem to have been manually updated and focused on written news texts, prioritizing journalistic autonomy. The BBC changed the volume of the text as the pandemic evolved; ERR updated its textual information in response to changing pandemic conditions and governmental regulations (e.g., advice on wearing masks). The SVT and NRK DVPs were similar in having more automated updating that required less authorial intervention.

These observations suggest that the automation of specific processes—seemingly a matter of simple resource allocation—directly impacts the outcome. Therefore, the balance of autonomy, automation and time as well as the intertextual link between visualization, multimodal content and text are essential editorial decisions that ultimately affect the communicative function. Additionally, changes in focus and the choice of news values partly affect the automation of processes, as they may hinder or support these editorial choices and influence the communicative function.

5.3. Communicative function

The digital form of information presentation embraces several categories that foster or hinder the use and understanding of visualization. Categories such data literacy, visual framing, storytelling and engagement make up the communicative function domain. Graphs, tables, and charts must explain what is presented, and users require guidance on the principles by which the data have been chosen, collected and framed. This increases transparency and credibility.

5.3.1. Data literacy

When graphs are used, the audience is expected to know how to interpret the types of chart and graph that present the data. Users, especially those with limited data literacy, may be misled by the wrong type of graph (or manipulation of the visualization), rare or specific graph types or data presentations whose interpretation demands mathematical and statistical sophistication. Most of the graphs and charts in our analysis presented limited data in a few types of graph, which were largely

comprehensible to people with lower levels of data literacy. However, data activist DVPs presented data without framing, creating the potential for multiple (mis)interpretations. Approaching the graphs of our DVP samples from the perspective of Potter's (2004) *interpretation* and *shared meaning* axioms, we found Potter's conditions to be mostly fulfilled. However, this was true only narrowly of individual graphs. The whole DVPs with various types of graph left significant room for interpretation. The DVPs also had shortcomings related to Potter's *power* and *purpose* axioms, generally presenting little or no information on the media functionality that would empower users to benefit from the presented knowledge. Similarly, the purpose axiom was unfulfilled in most of the DVPs (except those of ERR and the BBC), as the users were left to figure out how to shift control from the media to themselves.

Users may expect sources to be transparent for verification purposes, but efforts at transparency may backfire by creating the 'illusion of transparency'. This became evident in Worldometer's data activist DVP, which aggregated numerous databases of uneven quality and reliability. The data sources also changed over time. The links and references to the sources seemed to provide transparency but instead revealed patchy data and outdated information, thus increasing distrust of the whole data visualization, as misleading data and untrustworthy sources reduce credibility. Data and other information sources need to be regularly monitored for timeliness, even if the visualization is automatically updated. Our analysis suggests that the creators of data visualizations should assess the target audience's data literacy or add complementary material to support the learning of new skills and the gaining of knowledge from the DVP.

5.3.2. Visual framing

Deciding how to tell the story begins with selecting news values and focus in the editorial layer. This manifests in the visual framing through colors, figures, symbols and graphs and inspires emotions, opinions, facts, education and entertainment. In the first round of collection (May 15, 2020), our selected DVPs highlighted morbidity, danger and negativity. Visual framing in DVPs may combine several elements, but, in our sample, the main element for characterizing the message was color. Hence, we focus on the use of color in the sample DPVs.

Most of the data visualizations used a stereotypical color palette—black for deaths, red for new cases, green for the number of recovered people. The exceptions were Koroonakaart.ee and NRK, which used pastel colors in various shades of blue and green. In the second round of data collection (July 15, 2020), most of the visualizations had changed their colors to less dramatic ones or softened the colors' tones. This shows that the creators recognized how to dramatize the topic through visual framing. From the journalistic point of view, this sort of accentuation helps to deliver the news message. Nevertheless, in the context of continually evolving events and their data, visualizations may inspire panic by sensationalizing morbidity or may quickly induce fatigue as the audience becomes inured to over-dramatization. As in

the fairy tale of the boy who cried wolf, the 'first wave' coronavirus visualizations tended to over-dramatize the situation while the statistically very similar situation in the autumn inspired much less dramatic visualizations.

Data visualizations have not developed conventional genres or other format-related attributes that guide and support recipients in understanding the message. Stories with more text than graphs can employ classic news or feature genres, but relying on traditional narrative types does not mean that the text must dominate over visualizations. In our selection, SVT and NRK (as examples of PSM) and JHU (representing data activist DVPs) stood out by reason of their strong intertextuality between the visualization and the text.

Nevertheless, these DVPs' fragmented nature presented multiple focal points and did not send one clear message. In comparison, the BBC and ERR combined conventional news and feature narratives that focused more on the text's central message. However, the text was rather loosely related to the data visualization. Although the BBC and ERR communicated a more transparent and focused message, the intertextuality between text and visualization was modest and obfuscated the communicative function of the data visualization. All the PSM data visualizations had the potential to employ scrollytelling (i.e., engaging users through the interactive navigation of media), but none fully exploited this technique.

The data activist DVPs Koroonakaart.ee and Worldometer as well as the public authorities' DVPs offered visualizations that could be described in the phrase 'do what you want with the data', as the creation and interpretation of the narrative and message were left to the user. These DVPs were presented without much explanation or any detectable story type. The examined data activist DVPs seemed intentionally to avoid structuring. The graphs often included the technical asset of an embedding code for presentation on other web pages. Hence, the user stood in the midst of various data piles requiring analysis and interpretation. Personal observations indicate that such open data sets were simultaneously used for doomsday predictions and arguments that there was nothing to worry about.

The presentation of charts and graphs without explanatory information enables the creation of new and distorted contexts and interpretations of the data, suggesting a need to consider the interpretation and shared meaning axioms of media literacy. In this light, the activists' initiatives of presenting the data for users to interact with as they pleased led to potentially hazardous examples of mis-, mal- and disinformation. It is crucial to achieve a balanced intertextuality between the visualization and text, which strengthens the communicative function.

5.3.3. Engagement

We divided engagement with the data visualization into practical and technical usability, as the engagement takes place on two levels. The level of *practical usability* is by its nature more closely related to content (the substantive element) and involves the information's relevance to the user. *Technical usability* describes the means of

technical interactivity that influence the user experience. The analysis of the DVPs shows that authors attempted to achieve practical usability through the news values emphasized in the visualization, which also affected the possible emotional engagement with the visualization. The DVPs presented the number of deaths and infected cases at the top of the page, prioritizing *morbidity* and the *frequency* that is important to show the evolution of the event. Later, the mainly journalistic DVPs added regional information that emphasized *geographic proximity*. The analysis of engagement at the technical usability level shows that there were sometimes explanations on the graphs of what a particular symbol or element meant (primarily on the PSM DVPs). However, except for those of NRK and SVT, all the DVPs lacked instructions on the technical use of interactive visualizations, which are essential for less digitally advanced users.

6. A HEURISTIC MODEL OF CREATING A DATA VISUALIZATION

We integrated various concepts, previous studies and the results of the iterative coding process in devising our heuristic model, an analytical framework that enables journalists and data visualizers to better understand and create DVPs that consider the user perspective. The model presented in Figure 4 summarizes the three domains affecting the creation of the DVP.

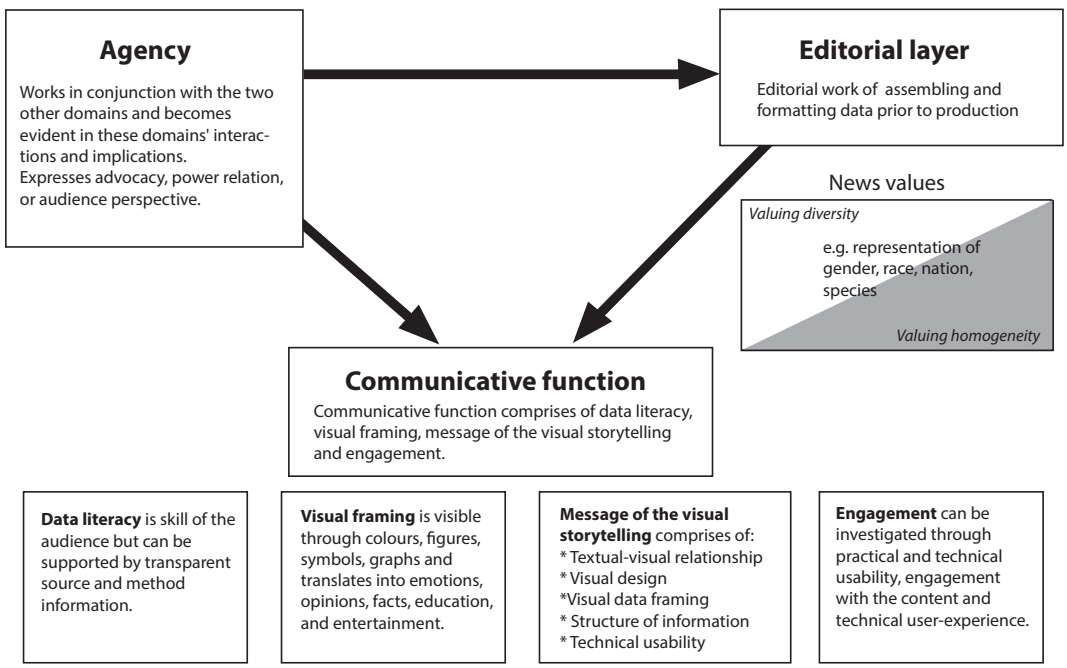

Figure 4. Interrelation of domains in the heuristic model of continually evolving DVPs

In Figure 4, the *agency* domain represents the element related to the dimensions of advocacy, power relations and audience perspective. Agency works in conjunction with the two other domains and becomes evident in those domains' interactions and implications.

The *editorial layer* domain determines the production process, which strongly influences the outcomes of the communicative function domain by focusing on specific news values. The editorial layer also includes decisions about the level of automation of routine processes. In the case of continuous coverage, automation decreases the workload and human errors and frees time that the DVP's authors can use for more intellectually stimulating tasks. The message is framed on one side by the communicative function and on the other by the news values emphasized in the narrative. In this sense, news values act as the criteria that the author(s) highlight in addressing the user; this attribute is relevant not only to journalists but also to data activists and other data storytellers.

The *communicative function* domain embraces the tacit knowledge, skills and competencies for understanding the presented information and judging its credibility. It reflects the message of the visualization and storytelling, which embraces aspects of the visual framing of the data, engagement, technical usability and the storytelling aspects that influence the use of the data and the reception of the message. Knowledge, skills and competencies related to data literacy are characteristics of users, but, in this mode, we use them to indicate what a DVP's creators must consider and how they can support the audience. They describe what the DVP requires from the user and the extent of support offered to interpret the data (e.g., its limitations, biases, frames and the transparency of its sources). Visual framing and storytelling embrace elements related to understanding the visual presentation of the data (e.g., the textual-visual relationship, framing the data with visual signifiers and the structure in which the information is presented). As the data can be visualized at several levels of complexity, it is relevant to assess the visualization's technical usability (e.g., technical guidelines and explanations for the user).

While the *communicative function* domain deals with aspects that are more dependent on the audience's side, the *agency* and *editorial layer* domains involve newsroom practices, which is why 'automation' is an essential subcategory domain.

7. CONCLUSIONS

This article deals with how events that evolve over time are expressed in DVPs that eventually create the event's image for an audience. Covid-19 inspired (and continues to inspire) various global and national journalistic DVPs of the pandemic's development. Beyond the pandemic, similar reporting may emerge in wars or catastrophes that evolve over time. Our analysis shows (1) how variations in diverse actors' agency explain their position and role-taking, (2) how the editorial layer reflects the balance

between automation, time budgets and author autonomy and (3) that the overall communicative function of the DVP supports the receiver in interpreting the visualization. The heuristic model developed in this study can be used by practitioners, e.g. reporters and editors in newsrooms, government officials, data activists, NGOs and others who create continually evolving DVPs to cover datafied events. As a tool, the model would be most beneficial as a guideline in the preliminary stages of planning and developing a DVP. It would also serve well as a framework in further research on data visualization practices (e.g., in comparing newsrooms' and data activists' initiatives).

The author's *agency* plays a crucial role but may often be unrecognized in the creative process. Public authorities perform the roles of service providers and loyal facilitators of power while data activists, who claim to act as neutral disseminators, also take an advocacy or watchdog role. Journalists who mediate the data visualizations of public authorities and data activists also assume the actor roles of the visualizations' initial authors. Our results indicate that it is crucial to discuss agency in the initial stages of creating a DVP, as agency seldom changes during the process. Future studies can potentially ask how agency and (journalistic) roles are negotiated in the work processes.

Agency's function cannot be undermined, as it partly determines the angle, focus and the selection of the story's news values as well as the *editorial layer*. The analysis of the editorial layer shows that the selection of news values directly influences how the data are presented (e.g., emphasizing geographical or emotional proximity, conflict or morbidity). Choices related to the representation of gender, race, nationality or even species are also based on the editorial layer. The editorial layer determines the balance between automating updates and the author's autonomy to change the visualization's message, and this balance influences the variety of news values presented in the story. While balancing autonomy and resources may be a familiar challenge in newsrooms, data activist initiatives and public authorities are only now discovering it. The data activists' and public authorities' DVPs in our sample indicate that they may create data visualizations for third parties without fully considering the media literacy aspect, which leads to potential misinterpretation of the data in a false context. We underscore that data visualizations are today often digital products that require usability testing, like any other digital service, especially if created by agents who have no previous professional experience in creating content for the public.

In analyzing the communicative functions of DVPs, we see that practical and technical usability are crucial factors that support the receiver in engaging with and interpreting the visualizations. The key to the clear, focused communicative function of a DVP is a balanced intertextuality between the visualization and text. If this link is weak, it obfuscates the message. Paying attention to the potential pitfalls in the communicative function must be the responsibility of the data visualizer.

Taking into consideration the categories of the three domains, the authors of any

DVP put themselves in the position of the user, and this enables them to detect and rectify potential shortcomings of the DVP as well as take decisions in the production process that make the DVP sustainable in the long term.

Dr. Marju Himma-Kadakas is a postdoctoral researcher at Karlstad University, where her research is part of the Ander Centre for Research on News and Opinion in the Digital Era (NODE). She is also a part-time Research Fellow of journalism studies at the University of Tartu, Estonia. She is the Estonian PI at the international project Journalistic Role Performance. Her research combines studies on news work, journalistic skills, and engagement of young audiences.

Prof. Pille Pruulmann-Vengerfeldt is a professor in media and communication, Malmö University. Her research interests have focused around datafied audiences, cultural citizenship and participation and engagement in museums, libraries and public broadcasting. She has also worked internet users and social applications of new technologies. She has been an active member in the European Communication Research and Education Association (ECREA) and is currently the international director of European Media and Communication Doctoral Summer School. She has published over hundred articles both in journals and as book chapters and has been in the team of editors for more than ten books. She is currently working within the Data Society research program focusing on the questions of digital culture and digital museums.

Dr. Signe Ivask is a postdoctoral researcher at Masaryk University, Faculty of Social Studies. Her research is supported by Postdoc2MUNI, CZ.02.2.69/0.0/0.0/18_053/0016952. She is also a research fellow at the University of Tartu, Estonia. Her research interests include stress and burnout of journalists, working routines in the newsroom, decision-making in newswork and communication between journalists and audience.

Acknowledgements
The authors would like to thank Åsa Harvard Maare, senior lecturer in graphic design at Malmö University, for her constructive feedback and dedicated work on visualizations of this article. Signe Ivask's research is supported by Postdoc2MUNI, CZ.02.2.69/0.0/0.0/18_053/0016952.

This research was partly supported by the Ander Foundation: Anne Marie och Gustav Anders Stiftelse för mediaforskning.

REFERENCES

Appelgren, E. (2018). An Illusion of Interactivity. *Journalism Practice, 12*(3), 308–325. https://doi.org/10.1080/17512786.2017.1299032

Carlson, M. (2015). The Robotic Reporter. *Digital Journalism, 3*(3), 416–431. https://doi.org/10.1080/21670811.2014.976412

Caswell, D., & Dörr, K. (2018). Automated Journalism 2.0: Event-driven narratives. *Journalism Practice, 12*(4), 477–496. https://doi.org/10.1080/17512786.2017.1320773

Coddington, M. (2015). Clarifying Journalism's Quantitative Turn. *Digital Journalism, 3*(3), 331–348. https://doi.org/10.1080/21670811.2014.976400

Engebretsen, M., Kennedy, H., & Weber, W. (2018). Data Visualization in Scandinavian Newsrooms. *Nordicom Review, 39*(2), 3–18. https://doi.org/10.2478/nor-2018-0007

Galtung, J., & Ruge, M. H. (1965). The Structure of Foreign News: The Presentation of the Congo, Cuba and Cyprus Crises in Four Norwegian Newspapers. *Journal of Peace Research, 2*(1), 64–90. https://doi.org/10.1177/002234336500200104

Gray, J., Gerlitz, C., & Bounegru, L. (2018). Data infrastructure literacy. *Big Data & Society, 5*(2), 2053951718786316. https://doi.org/10.1177/2053951718786316

Hanitzsch, T., Hanusch, F., Mellado, C., Anikina, M., Berganza, R., Cangoz, I., Coman, M., Hamada, B., Hernández, M. E., Karadjov, C. D., Moreira, S. V., Mwesige, P. G., Plaisance, P. L., Reich, Z., Seethaler, J., Skewes, E. A., Noor, D. V., & Yuen, E. K. W. (2011). Mapping Journalism Cultures Across Nations. *Journalism Studies, 12*(3), 273–293. https://doi.org/10.1080/1461670X.2010.512502

Hanitzsch, T., Ramaprasad, J., Arroyave, J., Berganza, R., Hermans, L., Hovden, J. F., Láb, F., Lauerer, C., Tejkalová, A., & Vos, T. P. (2019). 5. Perceived Influences: Journalists' Awareness of Pressures on Their Work. In *Worlds of Journalism* (pp. 103–132). Columbia University Press. https://www.degruyter.com/document/doi/10.7312/hani18642-006/html

Hanitzsch, T., & Mellado, C. (2011). What Shapes the News around the World? How Journalists in Eighteen Countries Perceive Influences on Their Work. *The International Journal of Press/Politics, 16*(3), 404–426. https://doi.org/10.1177/1940161211407334

Hanitzsch, T., & Vos, T. P. (2018). Journalism beyond democracy: A new look into journalistic roles in political and everyday life. *Journalism, 19*(2), 146–164. https://doi.org/10.1177/1464884916673386

Harcup, T., & O'Neill, D. (2001). What Is News? Galtung and Ruge revisited. *Journalism Studies, 2*(2), 261–280. https://doi.org/10.1080/14616700118449

Harcup, T., & O'Neill, D. (2017). What is News?: News values revisited (again). *Journalism Studies, 18*(12), 1470–1488. https://doi.org/10.1080/1461670X.2016.1150193

Hullman, J., & Diakopoulos, N. (2011). Visualization Rhetoric: Framing Effects in Narrative Visualization. *IEEE Transactions on Visualization and Computer Graphics, 17*(12), 2231–2240. https://doi.org/10.1109/TVCG.2011.255

Karlsson, M., & Clerwall, C. (2018). Transparency to the Rescue?: Evaluating citizens' views on transparency tools in journalism. *Journalism Studies, 19*(13), 1923–1933. https://doi.org/10.1080/1461670X.2018.1492882

Kennedy, H., Hill, R. L., Aiello, G., & Allen, W. (2016). The work that visualisation conventions do. *Information, Communication & Society, 19*(6), 715–735. https://doi.org/10.1080/1369118X.2016.1153126

Kennedy, H., Hill, R. L., Allen, W., & Kirk, A. (2016). Engaging with (big) data visualizations: Factors that affect engagement and resulting new definitions of effectiveness. *First Monday*. https://doi.org/10.5210/fm.v21i11.6389

Kõuts-Klemm, R. (2019). Data literacy among journalists: A skills-assessment based approach. *Central European Journal of Communication, 12*(3), 299–315. https://doi.org/10.19195/1899-5101.12.3(24).2

Masullo, G. M., Jennings, J., & Stroud, N. J. (2021). "Crisis Coverage Gap": The Divide between Public Interest and Local News' Facebook Posts about COVID-19 in the United States. *Digital Journalism, 0*(0), 1–22. https://doi.org/10.1080/21670811.2021.1965488

Mellado, C. (2015). Professional Roles in News Content. *Journalism Studies, 16*(4), 596–614. https://doi.org/10.1080/1461670X.2014.922276

Miles, M. B., & Huberman, A. M. (1994). *Qualitative Data Analysis: An Expanded Sourcebook*. SAGE.

Mooney, P., & Juhász, L. (2020). Mapping COVID-19: How web-based maps contribute to the infodemic. *Dialogues in Human Geography, 10*(2), 265–270. https://doi.org/10.1177/2043820620934926

Neale, J. (2016). Iterative categorization (IC): A systematic technique for analysing qualitative data. *Addiction (Abingdon, England), 111*(6), 1096–1106. https://doi.org/10.1111/add.13314

Örnebring, H., & Conill, R. F. (2016). 14 Outsourcing Newswork. In T. Witschge, C. W. Anderson, D. Domingo, & A. Hermida, *The SAGE Handbook of Digital Journalism* (pp. 207–221). SAGE Publications Ltd. https://doi.org/10.4135/9781473957909.n14

Pászto, V., Burian, J., & Macků, K. (2020). COVID-19 data sources: Evaluation of map applications and analysis of behavior changes in Europe's population. *Geografie, 125*(2), 171–209. https://doi.org/10.37040/geografie2020125020171

Pelizza, A. (2020). "No disease for the others": How COVID-19 data can enact new and old alterities. *Big Data & Society, 7*(2), 2053951720942542. https://doi.org/10.1177/2053951720942542

Pentzold, C., Fechner, D. J., & Zuber, C. (2021). "Flatten the Curve": Data-Driven Projections and the Journalistic Brokering of Knowledge during the COVID-19 Crisis. *Digital Journalism, 0*(0), 1–24. https://doi.org/10.1080/21670811.2021.1950018

Phillips, A. (2010). Transparency and the New Ethics of Journalism. *Journalism Practice, 4*(3), 373–382. https://doi.org/10.1080/17512781003642972

Potter, W. (2004). *Theory of Media Literacy: A Cognitive Approach*. SAGE Publications, Inc. https://doi.org/10.4135/9781483328881

Primo, A., & Zago, G. (2015). Who And What Do Journalism? *Digital Journalism, 3*(1), 38–52. https://doi.org/10.1080/21670811.2014.927987

Pruulmann-Vengerfeldt, P. (2020). The Ways of Knowing the Pandemic With the Help of Prompted Autoethnography. *Qualitative Inquiry*, 1077800420960133. https://doi.org/10.1177/1077800420960133

Rupar, V. (2006). How did you find that out? Transparency of the newsgathering process and the meaning of news. *Journalism Studies, 7*(1), 127–143. https://doi.org/10.1080/14616700500450426

Srivastava, P., & Hopwood, N. (2009). A Practical Iterative Framework for Qualitative Data Analysis. *International Journal of Qualitative Methods, 8*(1), 76–84. https://doi.org/10.1177/160940690900800107

Tong, J., & Zuo, L. (2021). The Inapplicability of Objectivity: Understanding the Work of Data Journalism. *Journalism Practice, 15*(2), 153–169. https://doi.org/10.1080/17512786.2019.1698974

Tuchman, G. (1973). Making News by Doing Work: Routinizing the Unexpected. *American Journal of Sociology, 79*(1), 110–131.

Weber, W. (2020). Exploring narrativity in data visualization in journalism. In M. Engebretsen & H. Kennedy (Eds.), *Data Visualization in Society* (pp. 295–312). Amsterdam University Press. https://doi.org/10.2307/j.ctvzgb8c7.24

Wu, S., Tandoc, E. C., & Salmon, C. T. (2019a). Journalism Reconfigured. *Journalism Studies, 20*(10), 1440–1457. https://doi.org/10.1080/1461670X.2018.1521299

Wu, S., Tandoc, E. C., & Salmon, C. T. (2019b). When Journalism and Automation Intersect: Assessing the Influence of the Technological Field on Contemporary Newsrooms. *Journalism Practice, 13*(10), 1238–1254. https://doi.org/10.1080/17512786.2019.1585198